

WHAT IS CLAIMED IS:

1. A speech processing system adapted to receive an input related to one of speech and text and process the input to provide an output related to one of speech and text, the speech processing system accessing a module derived from a phone set having a plurality of phones for a tonal language, the phones being used to model syllables used in the module, the syllables having an initial and final part, wherein the final part comprises a plurality of phones that jointly and implicitly carry the tonal information.
2. The speech processing system of claim 1 wherein each phone of the final part includes information about the tone.
3. The speech processing system of claim 1 wherein the tonal language comprises a plurality of different tones with different levels of pitch.
4. The speech processing system of claim 3 wherein the different levels of pitch comprise two categorical levels, and wherein each phone has a categorical level associated with it.
5. The speech processing system of claim 3 wherein the different levels of pitch comprise three categorical levels, and wherein each phone has a categorical level associated with it.

6. The speech processing system of claim 3 wherein the different levels of pitch comprise five categorical levels, and wherein each phone has a categorical level associated with it.

7. The speech processing system of claim 1 wherein each syllable comprises the same form having the initial and the final, the final having two phones carrying partial tonal information each.

8. The speech processing system of claim 1 wherein at least some of the syllables of the tonal language include a glide, the glide being embodied in the initial.

9. The speech processing system of claim 1 wherein the speech processing system comprises one of a speech recognition system and a text-to-speech converter.

10. The speech processing system of claim 9 wherein each phone of the final part includes partial information about the tone.

11. The speech processing system of claim 9 wherein the tonal language comprises a plurality of different tones with different levels of pitch.

12. The speech processing system of claim 11 wherein the different levels of pitch comprise two

categorical levels, and wherein each phone has a categorical level associated with it.

13. The speech processing system of claim 11 wherein the different levels of pitch comprise three categorical levels, and wherein each phone has a categorical level associated with it.

14. The speech processing system of claim 11 wherein the different levels of pitch comprise five categorical levels, and wherein each phone has a categorical level associated with it.

15. The speech processing system of claim 9 wherein each syllable comprises the same form having the initial and the final, the final having two phones carrying partial tonal information.

16. The speech processing system of claim 9 wherein at least some of the syllables of the tonal language include a glide, the glide being embodied in the initial.

17. The speech processing system of claim 16 wherein the tonal language comprises Chinese or a dialect thereof, such as Cantonese.

18. The speech processing system of claim 16 wherein the tonal language comprises Thai or a tonal dialect thereof.

19. The speech processing system of claim 16 wherein the tonal language comprises Vietnamese or a tonal dialect thereof.

20. A speech processing system adapted to receive an input related to one of speech and text and process the input to perform one of speech recognition and text-to-speech conversion in order to provide an output related to one of speech and text, the speech processing system accessing a module derived from a phone set having a plurality of phones for a tonal language comprising a plurality of different tones with different levels of pitch, the phones being used to model syllables used in the module, at least some of the syllables having an initial and final part, wherein a first set of the plurality of phones are used to describe glide dependent initials, and a second set of the plurality of phones are used to describe the final part, wherein the final part comprises a plurality of phones, each phone including partial tonal information.

21. The speech processing system of claim 20 wherein the different levels of pitch comprise two

categorical levels, and wherein each phone has a categorical level associated with it.

22. The speech processing system of claim 20 wherein the different levels of pitch comprise three categorical levels, and wherein each phone has a categorical level associated with it.

23. The speech processing system of claim 20 wherein the different levels of pitch comprise five categorical levels, and wherein each phone has a categorical level associated with it.

24. The speech processing system of claim 20 wherein at least one syllable comprises only the final having two phones carrying partial tonal information each.

25. The speech processing system of claim 20 wherein each syllable comprises the same form having the initial and the final, the final having two phones carrying partial tonal information each.

26. The speech processing system of claim 20 wherein the tonal language comprises Chinese or a dialect thereof such as Cantonese.

27. The speech processing system of claim 20 wherein the tonal language comprises Thai or a tonal dialect thereof.

28. The speech processing system of claim 20 wherein the tonal language comprises Vietnamese or a tonal dialect thereof.

29. A method of speech processing comprising:  
accessing a module having a phone set comprising  
a plurality of phones for a tonal language,  
the phones being used to model syllables,  
the syllables having an initial and final  
part, wherein the final part comprises a  
plurality of phones that jointly and  
implicitly carry the tonal information;  
utilizing the phone set to identify syllables  
corresponding to the input for performing  
one of speech recognition and text-to-  
speech conversion; and  
providing an output corresponding to one of  
speech recognition and text-to-speech  
conversion.